REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
MAY 2013	CONFERENCE PAPER (Post Print)	JUN 2010 – FEB 2011

4. TITLE AND SUBTITLE
ADAPTIVE MODULATION APPROACH FOR ROBUST MPEG-4 AAC ENCODED AUDIO TRANSMISSION

5a. CONTRACT NUMBER IN HOUSE
5b. GRANT NUMBER FA8750-11-1-0048
5c. PROGRAM ELEMENT NUMBER 62702F

6. AUTHOR(S)
D. Rawat, S. Kumar, S. Nagaraj, and J. Matyjas

5d. PROJECT NUMBER ANCL
5e. TASK NUMBER 62
5f. WORK UNIT NUMBER 07

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
San Diego State University
Department of Electrical and Computer Engineering
5500 Campanile Drive
San Diego, CA 92182

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory/Information Directorate
Rome Research Site/RITE
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RI

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-RI-RS-TP-2013-025

12. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. PA Case Number: 88ABW-2011-2495
DATE CLEARED: 3 MAY 2011

13. SUPPLEMENTARY NOTES
© 2011 IEEE. Proceedings IEEE Military Communications Conference. Baltimore, MD. 7-10, Nov 2011 This work is copyrighted. One or more of the authors is a U.S. Government employee working within the scope of their Government job; therefore, the U.S. Government is joint owner of the work and has the right to copy, distribute, and use the work. All other rights are reserved by the copyright owner.

14. ABSTRACT
In this paper, we present a robust transmission scheme for MPEG-4 AAC encoded audio data over fading wireless channels using Unequal Error Protection based on adaptive modulation and forward error correcting (FEC) codes. The encoded audio data is divided into three error sensitivity categories and a suitable combination of FEC and modulation order is determined for each category. The results are compared with non-adaptive (single order) modulation with FEC for the same channel conditions. Perceptual Evaluation of Audio Quality (PEAQ)-Objective Difference Grade (ODG) score has been chosen to measure decoded audio quality. Simulation results are presented as PEAQ-ODG score vs. channel signal to noise ratio at a fixed channel bandwidth. The proposed scheme achieves significant performance gain over the non-adaptive error protection scheme.

15. SUBJECT TERMS
Bandwidth , Bit error rate , Bit rate , Fading , Forward error correction , Modulation , Wireless communication

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					MICHAEL J. MEDLEY
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** (*Include area code*)
U	U	U	UU	6	N/A

Adaptive Modulation Approach for Robust MPEG-4 AAC Encoded Audio Transmission

Deepak Rawat, Sunil Kumar, Santosh Nagaraj
Electrical and Computer Engineering
San Diego State University, San Diego, CA, USA
rawat@rohan.sdsu.edu, {skumar, snagaraj}@mail.sdsu.edu

John D. Matyjas
Air Force Research Laboratory
Rome, NY, USA
john.matyjas@rl.af.mil

Abstract - **In this paper, we present a robust transmission scheme for MPEG-4 AAC encoded audio data over fading wireless channels using Unequal Error Protection based on adaptive modulation and forward error correcting (FEC) codes. The encoded audio data is divided into three error sensitivity categories and a suitable combination of FEC and modulation order is determined for each category. The results are compared with non-adaptive (single order) modulation with FEC for the same channel conditions. Perceptual Evaluation of Audio Quality (PEAQ)-Objective Difference Grade (ODG) score has been chosen to measure decoded audio quality. Simulation results are presented as PEAQ-ODG score vs. channel signal to noise ratio at a fixed channel bandwidth. The proposed scheme achieves significant performance gain over the non-adaptive error protection scheme.**

Keywords: MPEG-4 AAC, adaptive modulation, wireless channels, forward error correction codes, error sensitivity categories, unequal error protection (UEP).

I. INTRODUCTION

Wireless communication is susceptible to channels errors in the transmitted data due to fading, shadowing etc. On the other hand, the compressed audio data is very sensitive to errors which may lead to serious distortions and decoder crashes. Therefore, the MPEG-4 Audio Coding standard provides an error protection tool (EPTOOL) for protection against channel errors [1]. It uses the cyclic redundancy check (CRC) for error detection and the systematic rate compatible punctured convolutional (SRCPC) codes for bit error correction.

However, the error protection introduces significant amount of overhead bits depending on the channel quality. This demands more wireless bandwidth, which is a major constraint in wireless channels. Since the loss in different parts of compressed audio bitstream contribute significantly different amounts of distortion to the decoded audio quality, we design an unequal error protection (UEP) scheme in this paper. We first group the compressed audio frame data in three error sensitivity categories (ESCs) based on their contribution to the decoded audio quality. Our UEP scheme uses a combination of adaptive modulation and SRCPC to achieve the best audio quality for the bandwidth-limited AWGN and Rayleigh fading channels. We measure the audio quality by using the 'perceptual evaluation of audio quality – objective difference grade' or PEAQ-ODG score. Since lower order modulation provides more error robustness at the cost of reduced throughput, adaptive modulation is an effective framework for selective protection of different audio data categories. The simulation results demonstrated that our adaptive modulation scheme with SRCPC achieves better audio quality for various channel SNRs.

In Section 2, we describe the categorization of audio data in different ESCs. In Section 3, we describe the effects of the choice of modulation order and FEC code rate on audio quality. In Section 4, we outline the proposed algorithm for optimizing audio transmission using adaptive modulation. Finally, in Section 5, we present our simulation results for additive white Gaussian noise (AWGN) and Rayleigh flat fading channels.

II. ERROR SENSITIVITY CATEGORIES (ESCs)

We use the Advanced Audio Codec (AAC) - Main profile with the single channel element (SCE) syntax, and transmit it using the audio data transport stream (ADTS) format. Single ADTS frame is divided into three logical ESCs as discussed below [2], [3]:

ESC1: ESC1 consist of the most critical information, where any bit error is likely to cause decoder crash. The data in ESC1 consist of fixed header (28bits), variable header (28bits), and some parts of data block, such as the channel Id, tag, global gain, individual channel stream (ICS) info, including the prediction and section (Huffman codebook and length) information. ICS info may have 11 or 15 bits depending upon long or short window sequence, respectively. Similarly, each section may have 7 or 9 bits, 4 bits for Huffman codebook to be used for the section and 3 or 5 bits for section length [4], [5]. The size of ESC1 varies depending upon frame and source rate. The ESC1 thus covers all the critical information bits including prediction (if any) and all section information.

ESC2: MPEG-4 standard specifies protecting the first 192 bits of raw data block for SCE [1], [3]. In most cases, these bits are

Approved for Public Release; Distribution Unlimited: 88ABW-2011-2495, 3 MAY 2011. This work was partially supported by the Department of Air Force under Contract # FA8750-08-1-0078. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

978-1-4673-0081-0/11/$26.00 ©2011 IEEE 561

covered by ESC1; sometimes ESC2 covers remaining bits. Also, [3], [6] and [7] suggest protecting scalefactors with higher priority than spectral coefficients. The scalefactors are coded using a Huffman scalefactor codebook and each code length varies from 1 to 19 bits. The number of scalefactors depends on number of sections in the frame as each section is assigned with one scalefactor. In this paper, the size of ESC2 is selected to be 200 bits to cover most of the scale factors, pulse, noise substitution and gain control (PNG) information. The ESC2 also contains important information to properly decode the spectral coefficients. Error in this part of bitstream, sometimes, cause decoder to stop decoding, as corruption may lead to invalid Huffman codes for scalefactors or wrong PNG information. Also, errors in scalefactors may cause distorted audio [3].

ESC3: The contents of ESC3 mostly consist of Huffman encoded spectral coefficients, and scalefactors and/or PNG if not covered by the previous category. Error(s) in spectral coefficients may lead to some moderate distortion in decoded audio quality, which can be concealed with error concealment schemes [3]. At a given target/source bitrate, the size of ESC3 varies from frame to frame and calculated as:

Size of ESC3= [AAC frame length (from variable header)]-[size of (ESC1 and ESC2)]

The AAC frame length varies from frame to frame depending upon type of input signal and target bitrate. Errors in ESC1 and ESC2 are referred to as syntax error, and errors in ESC3 are referred as data error [3].

III. SELECTION OF FEC AND MODULATION

Significant work has been done [2], [7] on transmission of UEP audio bitstreams with single modulation and non-uniform FEC. In this paper, we implement the UEP scheme using adaptive modulation with uniform FEC. The adaptive modulation takes advantage of the logical distribution of bits in an audio frame as mentioned in the previous section. The data throughput achieved at a given channel physical bandwidth (in kHz) increases when the modulation order is increased. However, the bit error rate (BER) also increases at the same time. Therefore, we have selected the modulation order based on the significance of each ESC: 4QAM for ESC1, 8QAM for ESC2 and 16QAM for ESC3.

Same FEC (SRCPC code) strength is selected for all ESCs to keep the algorithm complexity low as explained in the next section. Since the ESC1 contains only critical data and size of ESC2 is fixed at 200 bits (its contents may vary for ESC2), switching to a higher source rate would increase the data in ESC3. The ESC3 mainly contains the coded spectral coefficients, where each coefficient may vary from 1 to 16 bits depending upon the Huffman codebook used, which are less prone to channel errors. This ESC distribution and the use of adaptive modulation gives us flexibility to switch to a higher source rate at a given channel bandwidth, which is not possible using single (non-adaptive) modulation, such as 4-QAM for all ESCs.

A channel bandwidth of 64 kHz allows transmission of only 64 ksymbols per second. In case of BPSK modulation the symbol rate is equal to bit rate and 64kbps encoded audio data can be transmitted without FEC. In case of random channel errors requiring the use of FEC, the encoded source bit rate should be decreased to accommodate FEC parity bits, thus lowering the audio quality. We have used the perceptual evaluation of audio quality (PEAQ) tool to evaluate the audio quality under various channel conditions. PEAQ tool gives audio quality score as Objective difference grade (ODG) as shown in Table 1.

Table 1 specifies the perceptual interpretation of the ODG. Subjective Difference Grade (SDG) = Grade $_{\text{Signal under test}}$ − Grade $_{\text{Reference signal}}$ (from listening tests), where the Grade of reference signal is 5.00 [9]. The PEAQ algorithm correlates the PEAQ-ODG score to the SDG using human hearing and cognitive model [8], [9]. Freely available PEAQ basic model, “PQevalAudio,” is used in this paper which is available as a part of AFsp programs and subroutines from the Telecommunications & Signal Processing Laboratory at the University of McGil, Canada.

Table 1: PEAQ-ODG Score [6]

Impairment	ITU-R Five Grade Impairment Scale	SDG/PEAQ-ODG Score
Imperceptible	5.00	0.00
Perceptible, but not Annoying	4.00	-1.00
Slightly annoying	3.00	-2.00
Annoying	2.00	-3.00
Very annoying	1.00	-4.00

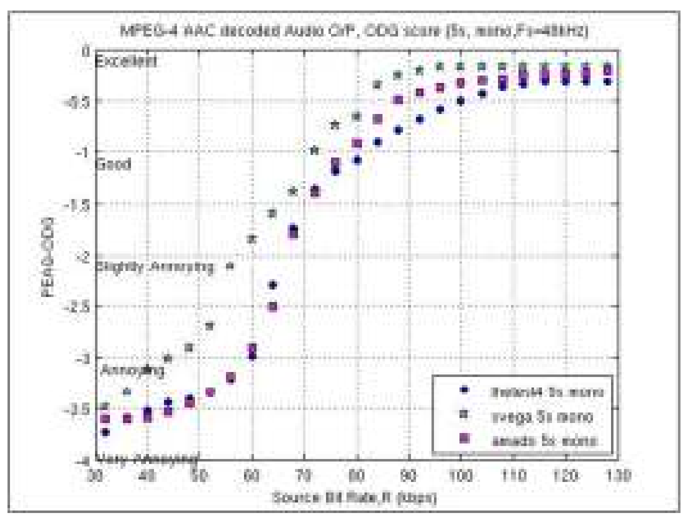

Figure1: PEAQ-ODG vs. Source bitrate for three audio signals.

Figure 1 shows the PEAQ-ODG score vs. source bitrate for three different audio streams. For example, at high channel Eb/No (11dB on additive white Gaussian noise (AWGN)

channels), BPSK will provide an average PEAQ-ODG score of -2.23 for the audio stream used in this paper which translates to perceptual audio quality in the range of *slightly annoying to annoying* as per the ITU-R BS.1387-1 standard [8], [9] shown in Table 1. But in case of QPSK/4-QAM, again at high Eb/No (negligible BER), the source rate can be switched to 128kbps (ignoring other transmission overhead) thus allowing transmission of high quality audio with PEAQ–ODG score of -0.4 which corresponds to *near excellent* audio quality. Here, Eb is the average energy per bit and No is the power spectral density of the Gaussian noise at the receiver.

The proposed scheme identifies the optimal choice of overall source rate and the FEC code rate for a given modulation order for transmission of each ESC. As the BER increases with decrease in channel Eb/No, the audio quality degrades gracefully by gradually lowering the source rate in order to accommodate FEC overhead for robust transmission. Our scheme takes advantage of the rrror tolerance of each ESC of ADTS audio frame and applies FEC and suitable modulation to each category resulting in higher perceived audio quality as compared to using the non-adaptive modulation. However, at low Eb/No, it is sometimes not possible to transmit audio without critical errors even with the lowest available code rates.

IV. SIMULATION SETUP AND PROPOSED ALGORITHM

As stated earlier, the non-adaptive scheme uses the 4-QAM modulation, whereas the adaptive modulation scheme employs the 4, 8, and 16 QAM for ESC1, ESC2 and ESC3, respectively. The FEC codes are applied using the error protection tool (EPTOOL) provided by the MPEG-4 v2 Audio standard [1]. The channel is simulated using the error generation (errGen) tool also provided in the same reference source code. The errGen tool requires BERs along with other parameters such as the source bit rate, seed and FEC protected bit stream. Channel specific BERs are derived using Matlab functions 'berawgn' and 'berfading'.

The EPTOOL allows applying SRCPC and/or CRC to each ESC. The mother code rate is ¼, thus, giving a vast range of code rates ranging from 8/8 to 8/32, where 8/8 provides no extra protection with no parity bits and 8/32 provides maximum protection with 3 parity bits per information bit. The SRCPC gives very good flexibility to adapt the rate according to the modulation order.

The algorithm described below (also shown in Figure 2) is used for adaptive as well as non-adaptive modulation strategies. The only difference is that for non-adaptive modulation all ESCs use single order modulation.

<u>*Algorithm:*</u>

1. Choose a value of Eb/No and channel bandwidth 'B'.
2. Start with the highest available source bit rate $R_{\max}$ and the highest FEC code rate $8/(8+a_{min})$ using the chosen modulation scheme, where 'a' is the SRCPC code rate for all ESCs. In the EPTOOL, the value of a varies from 0 to 24 corresponding to the code rates 8/8, 8/9,…, 8/32.
3. If all the audio frames can be received and decoded, compute the PEAQ - ODG score and go to step 5. Note that the PEAQ tool needs all the frames of the decoded file to calculate its ODG score by comparing with the original file. This algorithm tries all the possible combinations for the best ODG score along with meeting the minimum successful transmission criterion (i.e., minimum 4 successful attempts out of 5); this guarantees best ODG score at given Eb/No.
4. Otherwise, use a lower source rate to accommodate higher error protection, *i.e.*, new $R_{\max} = (R_{\max}\text{-}2N)$ and/or $a = a+1$, where N is a positive integer ≥1 (depending upon available physical channel bandwidth). The step size for source bit rate adaptation is 2kbps for gradual quality degradation. Go to step 2 with new values of $R_{\max}$ and $a_{\min}$.
5. Repeat steps 2 and 3 for five different seed values to generate different randomly corrupted bitstreams. This provision is present in MPEG-4 v2 Audio "errGen" tool.
6. Record the average PEAQ-ODG score for this Eb/No value after minimum four successful transmissions out of five attempts.
7. If there are no more source rate and SRCPC code rate combinations available for error protection, record the output as '*Unable to transmit without critical errors*' and assign PEAQ ODG score of -4.
8. Repeat for other values of Eb/No in the operational range.

AWGN and Rayleigh Fading channels are simulated with errGen tool with their respective BERs. The upper limit of the operating Eb/No range is selected such that with non-adaptive modulation (4-QAM in this case), it is possible to transmit near excellent audio quality with minimum protection. On the other side, the lowest Eb/No is selected such that minimum acceptable audio quality can be maintained with non-adaptive 4-QAM and a low rate FEC code. The lowest acceptable audio quality (PEAQ-ODG score of -3.5) and the highest achievable audio quality (PEAQ-ODG score of -0.3) correspond to source bit rates of 36 kbps and 96 kbps, respectively, for "thetest4.wav" audio clip of duration 5s and sampled at 48 kHz. This audio clip was also used in Figure 1.

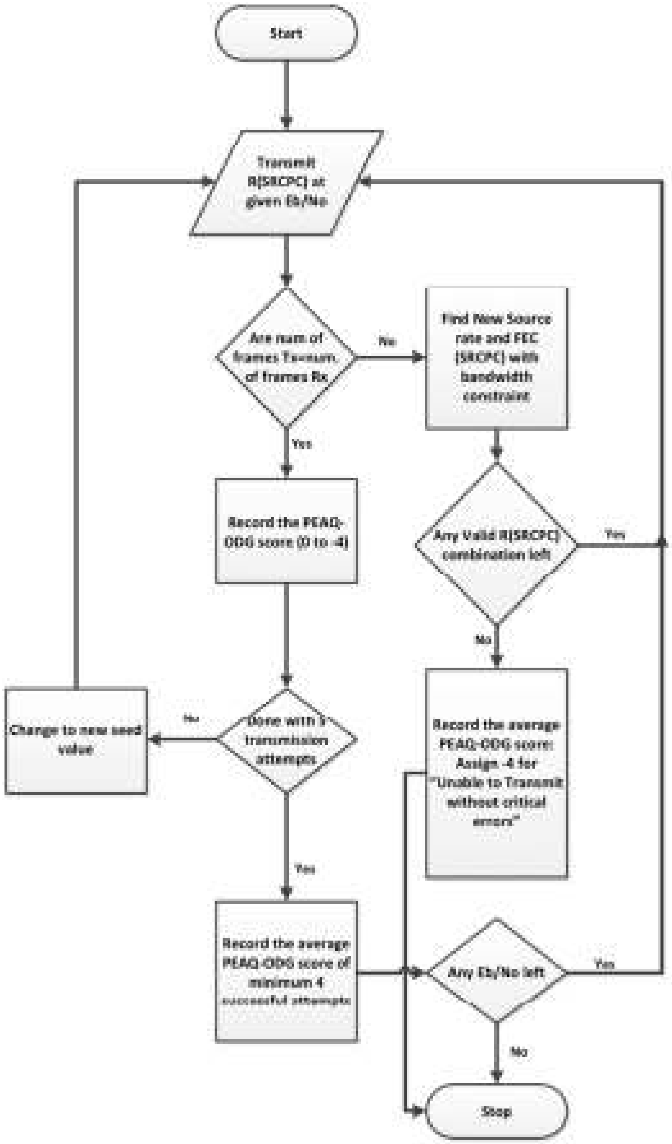

Figure 2: Flow chart of the proposed algorithm.

V. SIMULATION RESULTS

AAC encoded and SRCPC protected audio data were transmitted over bandwidth-limited AWGN and Rayleigh flat fading channels using the adaptive and non-adaptive modulation schemes as discussed in previous sections. We used two channel bandwidths of 64 kHz and 48 kHz for both channel models. The comparative performance of both schemes is discussed below.

A. AWGN Channel at 64kHz

As shown in Figure 3, the adaptive modulation provides better audio quality than the non-adaptive modulation scheme at Eb/No ≥ 9dB. At each Eb/No, the best audio source rate and SRCPC code rate is shown as 'R (code rate)'. However, the non-adaptive modulation scheme outperforms the adaptive modulation scheme at low Eb/No < 9dB due to high BER in ESC2 and ESC3. At Eb/No BER < 7dB, the data is corrupted by the very high BER and the adaptive modulation is unable to transmit without critical errors.

B. AWGN Channel at 48kHz

Results are similar to the 64 kHz channel case as shown in Figure 4. The adaptive modulation provides better audio quality than the non-adaptive modulation scheme at Eb/No ≥ 8dB. However, the non-adaptive modulation scheme outperforms the adaptive modulation scheme at low Eb/No < 8dB due to high BER in ESC2 and ESC3.

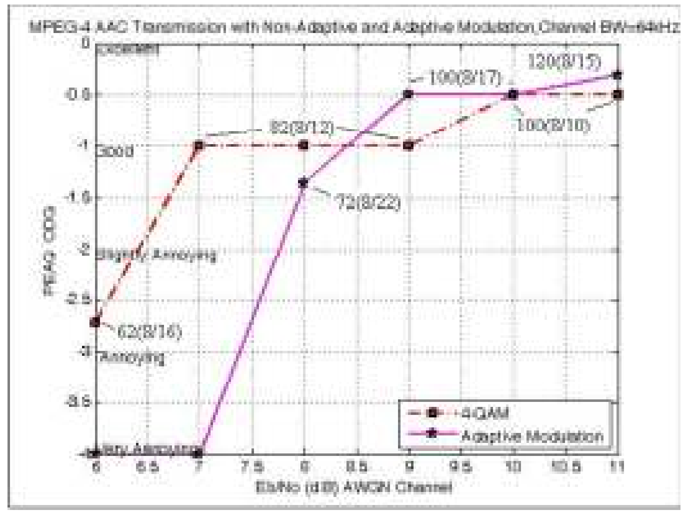

Figure 3: The PEAQ-ODG vs. Eb/No (dB) performance of both schemes for 64 kHz AWGN channel.

Therefore, even at 7dB, the transmission cannot achieve the desired success rate as compared to the AWGN-64kHz case.

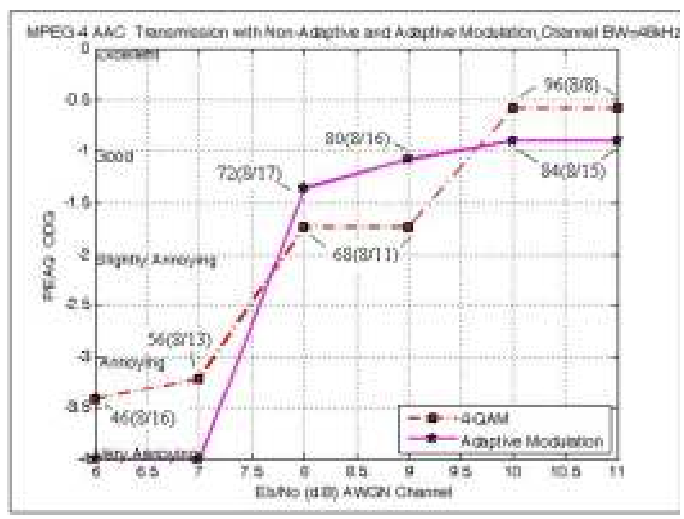

Figure 4: The PEAQ-ODG vs. Eb/No (dB) performance of both schemes for 48 kHz AWGN channel.

C. Rayleigh Fading Chanel at 64kHz

For Rayleigh fading channels, the operating Eb/No in our simulations ranges from 15 dB to 24 dB. As shown in Figure 5, the adaptive modulation scheme provides considerably better audio quality for these channel SNRs. The quality of both schemes is very poor for Eb/No < 15dB.

D. Rayleigh Fading Chanel at 48 kHz

Figure 6 shows that the adaptive modulation scheme provides significantly better audio quality than the non-adaptive modulation for Eb/No values above 15dB. At Eb/No < 16dB, the BER is too high, which makes it difficult to transmit without critical errors.

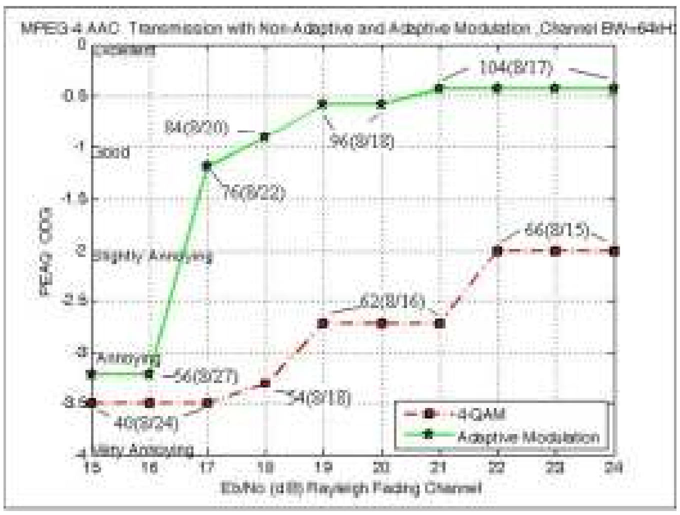

Figure 5: The PEAQ-ODG vs. Eb/No (dB) performance of both schemes for 64 kHz Rayleigh fading channel

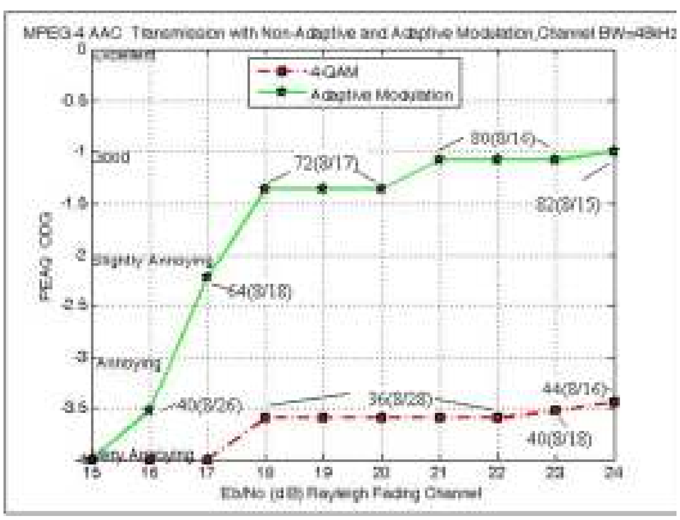

Figure 6: The PEAQ-ODG vs. Eb/No (dB) performance of both schemes for48 kHz Rayleigh fading channel.

VI. CONCLUSION

Bandwidth is a precious and expensive resource in wireless communication channels. Furthermore, the compressed audio data is very sensitive to errors in wireless channels. Since different parts of compressed audio bitstream, if lost, contribute significantly different amounts of distortion to the decoded audio quality, we designed an unequal error protection scheme in this paper. We divide the compressed audio frame data in three categories based on their importance to the decoded quality. Our UEP scheme uses adaptive modulation where modulation order is dependent on the bitstream importance. Our scheme determines the most suitable source bit rate and RCPC code rate to achieve the best audio quality on the bandwidth-limited AWGN and Rayleigh fading channels. The simulation results demonstrated that our adaptive modulation scheme with SRCPC achieves better quality for various channel SNRs.

REFERENCES

[1] ISO/IEC 14496-3 “Information Technology - Generic Coding of Audiovisual Objects,” Part 3: Audio, Subpart 4: General Audio (GA) Coding: AAC/TwinVQ, 2000.

[2] Y. Zhang, “Robust audio Coding over Wireless Channels,” IEEE WCNC, 2007.

[3] O. Derrien, M. Kieffer and P. Duhamel, “Joint Source/Channel Decoding of Scalefactors in MPEG-AAC Encoded Bitstreams,” in EUSIPCO, Switzerland, 2008.

[4] ISO/IEC 13818-7 2004(E) “Information Technology - Generic coding of moving pictures and associated audio information,” Part 7: Advanced Audio Coding (AAC), 2004.

[5] M. Bosi and R. E. Goldberg, “Introduction to Digital Audio Coding and Standards,” 2003, Kluwer Academic Publishers, Boston.

[6] R. Hu, X. Huang, M. Kieffer and P. Duhamel, “Robust Critical Data Recovery for MPEG-4 AAC encoded Bitstreams,” in IEEE ICASSP, 2010, pp. 397-400.

[7] A. Sieber and C. Weck, “What is the difference between DVB-H and DAB-in the mobile environment?” EBU (European Broadcast Union) Technical Review, July 2004.

[8] M. Šalovarda, I. Bolkovac and H. Domitrovic, “Estimating Perceptual Audio System Quality using PEAQ Algorithm,” in the 18th International Conf. Applied Electromagnetics and Communications, 2005.

[9] Recommendation ITU-R BS.1387-1, “Method for Objective Measurements of Perceived Audio Quality,” 1998-2001.